

Mechanism Design for Fair Allocation

Abhinav Sinha and Achilleas Anastasopoulos

EECS Department, University of Michigan

Email: {absi,anastas}@umich.edu

Abstract—Mechanism design for a social utility being the sum of agents’ utilities (SoU) is a well-studied problem. There are, however, a number of problems of theoretical and practical interest where a designer may have a different objective than maximization of the SoU. One motivation for this is the desire for more equitable allocation of resources among agents. A second, more subtle, motivation is the fact that a fairer allocation indirectly implies less variation in taxes which can be desirable in a situation where (implicit) individual agent budgetary constraints make payment of large taxes unrealistic. In this paper we study a family of social utilities that provide fair allocation (with SoU being subsumed as an extreme case) and derive conditions under which Bayesian and Dominant strategy implementation is possible. Furthermore, it is shown how a simple modification of the above mechanism can guarantee full Bayesian implementation. Through a numerical example it is shown that the proposed method can result in significant gains both in allocation fairness and tax reduction.

I. INTRODUCTION

Mechanism design is a well-established framework for dealing with decentralized resource allocation problems in the presence of strategic agents. The corresponding literature is vast, especially in the domain of Dominant, Nash and Bayesian implementation, [1], [2], [3], [4]. In the area of Engineering—and in particular in the area of Networks— a majority of the works start with the assumption that the social objective is the sum of individual utilities (SoU) of all the system’s agents.

There are strong mathematical reasons for preferring SoU as the resource allocation objective. The main one being that for SoU, in conjunction with quasi-linear utilities, agents’ individual goals can be aligned directly with the overall social objective, so that when agents maximize their net utility they are simultaneously maximizing the overall social objective. The VCG mechanism [5] is the most prominent example of this. With social objective $\sum_{i=1}^N v(\hat{x}(\theta); \theta_i)$ and individual utility of $v(\hat{x}(\phi); \theta_i) - t_i(\phi)$, VCG taxes $t_i(\phi) = -\sum_{j \neq i} v(\hat{x}(\phi); \phi_j) + f_i(\phi_{-i})$ precisely ask the user to perform social objective maximization whilst maximizing self utility.

There are, however, a number of problems of theoretical and practical interest where a designer may have a different objective than maximization of the SoU. One obvious reason for such a preference is the desire of the social planner to introduce fairness in the allocation process. Consider for example, a network where optimizing SoU results in one (or a few) agents receiving almost all the available resources and everyone else receiving an appreciably lower portion. In instances of this kind, appealing to fairness or equality, a

system designer may genuinely want allocations which are more equitable, even if this may come at the cost of reduced revenue.

A second—more subtle—reason for wanting a different social objective is the fact that the standard mechanism design framework does not provide any formal way of limiting the range of the taxes/subsidies required at equilibrium. This implies that when strong budget balance is imposed on the mechanism, the magnitude of the monetary transfers (taxes/subsidies) can vary greatly among agents (with those who benefit more from the allocation having to contribute more as well). This can be a significant practical problem since it does not take into account the budget constraints of individual agents. The work of [6] is attacking this problem by considering the dual version of the resource allocation problem and putting additional structure on the dual variables to ensure less variation between them. This would typically ensure less fluctuation in prices (and thus in the monetary transfers) since it is well-known that the dual variables for resource allocation optimization problems act as prices in the corresponding markets. A modified social objective provides an alternative way for dealing with the issue of large tax variation. By making the social objective a more concave function of the utilities (compared to the SoU) a smaller variation of the taxes is expected.

As soon as one moves away from the SoU objective, due to the social and individual objectives not aligning with each other, basic design techniques like VCG mechanism are not useful. This is one of the main reasons why there are significantly fewer results on fairness in the mechanism design literature. In [7], authors use the concept of “proportional fairness” (similar to the sum of log of utilities) to reduce disparity. The focus is on a tax-less mechanism where the contract proposes to throw away existing resources (“resource-burning”) in order to tax untruthful agents. This mechanism achieves at least $\frac{1}{e}$ fraction of proportionally fair allocation. In [8], optimal auctions in the Bayesian set up are derived, such that instead of efficiency maximization or revenue maximization, a linearly combined metric is maximized which favors exchange of goods at low prices (thereby ensuring fairer trade). “Envy-freeness” is another well-known criterion for equitable allocations (see [9]). This notion was originally proposed for exchange economies where an allocation is called envy-free if no agent is strictly better off by taking

someone else's allocation instead of their own¹. Such a notion was argued in terms of the stability it provides, since each agent may be content with what they have comparing to the possible option of acquiring someone else's allocation. This, however, is an ex-post notion and in general imposes quite stringent constraints on the design. For a large enough environment it may indeed be impossible to achieve envy-freeness in optimal allocation.

It is interesting to note that the problem of mechanism design for risk-averse agents (see for e.g., [10], [11]) is in some respects the opposite problem of the one addressed here. In that case, the social planner's objective is relatively more "aggressive" compared to the more risk averse individual objectives.

In this paper, we ask if and how we can design mechanisms that implement social objectives that are especially designed for fairness and go beyond the standard paradigm of SoU. We seek a methodology that is flexible enough to create space for the designer when envy-freeness may not be feasible. We concentrate on the form of the social objective given by the additive function $\sum_{i=1}^N g_\epsilon(v(x; \theta_i))$ where $v(x; \theta_i)$ is the utility of the i -th user with allocation x and type θ_i . Here we take $g_\epsilon(z) = z - \epsilon f(z)$ as a family of concave functions parameterized by $\epsilon > 0$, with $f(\cdot)$ an arbitrary convex function. In this setup, the SoU is a special case with $\epsilon = 0$, while as the parameter ϵ increases, more fairness is built into the allocation. The form $g_\epsilon(z) = z - \epsilon f(z)$ is considered without significant loss of generality, since it closely emulates Taylor's series (w.r.t. ϵ) of many interesting families of concave functions. Consider for example the family $h_\epsilon(z) = z^{1-\epsilon}$, the Taylor's series for this family is $h_\epsilon(z) = z - \epsilon(z \log(z)) + o(\epsilon)$, where $z \log(z)$ is indeed convex. Note that unlike agents' utilities, the choice of the specific function g_ϵ is in the designer's hand, as long as it serves the design objective of fairer allocation. Furthermore, in Section VII we discuss how the results are still valid even when $f(z)$ depends on ϵ (under certain conditions).

Within this framework we ask for which values of the parameter ϵ and under what conditions for the convex function $f(\cdot)$ is Dominant strategy implementation possible and when is Bayesian Nash Equilibrium (BNE) implementation possible. We show that indeed mechanism design is possible provided ϵ is not too large, by providing an upper bound on the range of ϵ . This is done by formulating the incentive compatibility constraints as a set of linear inequalities on the design variables and checking whether this system (together with strong budget balance and/or individual rationality) is feasible. Our proving techniques follow closely the work of [12], [13]. Not surprisingly, the results in the Bayesian set-up are derived under certain assumptions on the prior beliefs, $p_i(\theta_{-i}|\theta_i)$, of agents, which are trivially satisfied for the case where $p_i(\theta_{-i}|\theta_i) = p_i(\theta_{-i})$.

For the case of Bayesian mechanism design, we also propose a modification to our mechanism which ensures not only that truth-telling is a BNE but also that it is the only BNE. This modification doesn't make significant changes to

the mechanism and only requires exchange of one additional message.

We finally demonstrate (through a numerical example) that the range of ϵ is sufficient to provide quite significant gains in fairness - as measured by the decrease in Gini Coefficient of the utilities. In addition, and this relates to the second reason we mentioned above regarding our motivation for this work, the results show a significant decrease in the variance of required taxes to achieve incentive compatibility.

The remaining of this paper is organized as follows: Section II defines the Centralized problem which has been modified for fairness. The next two sections prove the existence of mechanisms that implement the aforementioned social objectives for type sets of size two in Dominant strategy (Section III) and for general type sets in BNE (Section IV). Section V presents the modification for full Bayesian implementation. The numerical example is presented in Section VI. Finally, Section VII discusses future work and immediate extensions of the results in here.

II. CENTRALIZED PROBLEM

For a system of agents $\mathcal{N} = \{1, \dots, N\}$, efficient allocation is calculated via the following optimization problem:

$$\hat{x}_\epsilon(\theta) = \arg \max_{x \in \mathcal{X}} \sum_{i \in \mathcal{N}} g_\epsilon(v(x; \theta_i)), \quad (1)$$

where $\mathcal{X} \subset \mathbb{R}_+^N$ is the constraint set, $\theta = (\theta_i)_{i \in \mathcal{N}} \in \Theta \triangleq \times_{i \in \mathcal{N}} \Theta_i$ is the type profile of agents and Θ_i is the discrete type set for agent i with $|\Theta_i| = L_i$. The utility function $v(x; \theta_i)$ measures agent's i satisfaction at allocation x with private type being θ_i . A concave transformation $g_\epsilon(\cdot)$ is applied for making the allocation fairer compared to the SoU setup. It is further assumed that the functions $g_\epsilon : \mathbb{R} \rightarrow \mathbb{R}$ and $v(\cdot; \theta_i) : \mathbb{R}_+^N \rightarrow \mathbb{R}$ are such that the optimization has a unique solution (e.g., if $v(\cdot; \theta_i)$ is concave and g_ϵ is concave and increasing and \mathcal{X} is a convex set). Specifically the form $g_\epsilon(z) = z - \epsilon f(z)$ for $\epsilon \geq 0$ is considered here, where $f(\cdot)$ is assumed to be a bounded convex function. Note that at $\epsilon = 0$ optimization (1) becomes the SoU problem, while as ϵ increases from 0 the function g_ϵ has a stronger concave component.

Define, $\forall i \in \mathcal{N}$, $\mathcal{F}_i := \{(\psi_i, \xi_i) \in \Theta_i^2 \mid \psi_i \neq \xi_i\}$. With this we define the difference, $\forall i \in \mathcal{N}$, $(\theta_i, \phi_i) \in \mathcal{F}_i$, $\theta_{-i} \in \Theta_{-i}$, $\epsilon > 0$,

$$K_i(\theta_i, \phi_i, \theta_{-i}, \epsilon) := \sum_{j \in \mathcal{N}} g_\epsilon(v(\hat{x}(\theta_i, \theta_{-i}); \theta_j)) - \sum_{j \in \mathcal{N}} g_\epsilon(v(\hat{x}(\phi_i, \theta_{-i}); \theta_j)). \quad (2)$$

Optimality conditions from (1) give that the above difference is always non-negative. However, due to the finite type spaces, the difference above is expected to be strictly positive. We make appropriate assumptions in this regard later.

¹Note that this exchange refers to both allocation of good and taxes.

III. DOMINANT STRATEGY IMPLEMENTATION

The Mechanism Design problem in this section is to find a message space $\mathcal{M} = \times_{i \in \mathcal{N}} \mathcal{M}_i$ and allocation, tax functions $(\tilde{x}, t) : \mathcal{M} \rightarrow \mathcal{X} \times \mathbb{R}^N$ such that the induced game for agents in \mathcal{N} with action space \mathcal{M} and quasi-linear utilities

$$\tilde{u}_i(m; \theta_i) = v(\tilde{x}(m); \theta_i) - t_i(m) \quad \forall m \in \mathcal{M}, i \in \mathcal{N} \quad (3)$$

has a dominant strategy equilibrium² m^* for which $\tilde{x}(m^*) = \hat{x}_\epsilon(\theta)$, where $\theta = (\theta_i)_{i \in \mathcal{N}}$ is the true type profile. This is known as Dominant Strategy Incentive Compatibility (DSIC).

In general, Dominant strategy implementation is very restrictive (note that the well studied VCG mechanisms are no longer applicable since this is not the maximization of SoU). Following [14], the special case of $L_i = 2 \forall i \in \mathcal{N}$ is considered in this section. In particular, $\Theta_i = \{\theta_i^H, \theta_i^L\} \forall i \in \mathcal{N}$.

The proposed mechanism is a direct mechanism, thus $\mathcal{M}_i = \Theta_i \forall i \in \mathcal{N}$. Agents report their types (possibly untruthfully) $\phi = (\phi_i)_{i \in \mathcal{N}}$ and the allocation they receive on the basis of this is the optimal allocation $\hat{x}_\epsilon(\phi)$ for the quoted type profile ϕ , where $\hat{x}_\epsilon(\cdot)$ is defined in (1).

Assuming that for not participating in the mechanism, an agent receives 0 utility value (including 0 tax), the voluntary participation condition for Dominant strategy implementation is $\forall (\theta_i, \theta_{-i}) \in \Theta, i \in \mathcal{N}$,

$$v(\hat{x}_\epsilon(\theta_i, \theta_{-i}); \theta_i) - t_i(\theta_i, \theta_{-i}) \geq 0. \quad (4)$$

This is the ex-post version of the individual rationality (IR).

The first contribution of this paper is summarized in the following Theorem.

Theorem III.1. *For any $\epsilon \geq 0$, there exist taxes $(t_i(\phi))_{\phi \in \Theta, i \in \mathcal{N}}$ that satisfy the corresponding DSIC (which implies implementation in Dominant strategies) and IR conditions if and only if, $\forall \theta_{-i} \in \Theta_{-i}, i \in \mathcal{N}$,*

$$v(\hat{x}_\epsilon(\theta_i^H, \theta_{-i}); \theta_i^H) + v(\hat{x}_\epsilon(\theta_i^L, \theta_{-i}); \theta_i^L) - v(\hat{x}_\epsilon(\theta_i^H, \theta_{-i}); \theta_i^L) - v(\hat{x}_\epsilon(\theta_i^L, \theta_{-i}); \theta_i^H) \geq 0. \quad (5)$$

Proof: Please see Appendix A. ■

Next we state the assumption on (1) under which the next result in this section is derived.

Condition (A_D): Assume that $\exists \epsilon_{max} > 0$ such that for all $0 \leq \epsilon < \epsilon_{max}$, $\forall i \in \mathcal{N}, (\theta_i, \phi_i) \in \mathcal{F}_i, \theta_{-i} \in \Theta_{-i}$,

$$K_i(\theta_i, \phi_i, \theta_{-i}, \epsilon) + K_i(\phi_i, \theta_i, \theta_{-i}, \epsilon) > 0. \quad (6)$$

Condition (A_D) and the ones below in Corollary III.3, IV.2, IV.3 and Condition (A_B) can be checked only at $\epsilon = 0$. By continuity of the optimization on parameter ϵ , this will imply that these conditions continue to hold for all $0 \leq \epsilon < \epsilon_{max}$ for some $\epsilon_{max} > 0$.

Theorem III.2. *If Condition (A_D) is satisfied then $\exists \tilde{\epsilon}_{max} > 0$ such that for all $0 \leq \epsilon < \tilde{\epsilon}_{max}$ there exist taxes $(t_i(\phi))_{\phi \in \Theta, i \in \mathcal{N}}$ that satisfy DSIC (which implies implementation in Dominant strategies) and IR.*

²For any θ , message $m^* \in \mathcal{M}$ is a dominant strategy equilibrium if it satisfies $\tilde{u}_i(m_i^*, m_{-i}; \theta_i) \geq \tilde{u}_i(m_i, m_{-i}; \theta_i) \quad \forall m_i \in \mathcal{M}_i, \forall m_{-i} \in \mathcal{M}_{-i}, \forall i \in \mathcal{N}$.

Remark. *The relation between ϵ_{max} and $\tilde{\epsilon}_{max}$ is defined in the proof, specifically via the relation in (27).*

Proof: Please see Appendix B. ■

Stated below is a corollary to the above result, which uses a stricter condition.

Corollary III.3. *If $\exists \epsilon_{max} > 0$ such that for all $0 \leq \epsilon < \epsilon_{max}$,*

$$\min_{i \in \mathcal{N}} \min_{(\theta_i, \phi_i) \in \mathcal{F}_i} \min_{\theta_{-i} \in \Theta_{-i}} K_i(\theta_i, \phi_i, \theta_{-i}, \epsilon) > 0, \quad (7)$$

then $\exists \tilde{\epsilon}_{max} > 0$ such that for all $0 \leq \epsilon < \tilde{\epsilon}_{max}$ there exist taxes that satisfy DSIC and IR.

Proof: As (7) implies (6), the Corollary follows from Theorem III.2. Also, as long as the above assumption holds, the value of $\tilde{\epsilon}_{max}$ will be same as the one derived in the previous proof, since it is only defined by the expression in (27). ■

IV. BAYESIAN IMPLEMENTATION

In this section the type sets are of arbitrary size. Dominant strategy implementation is too restrictive for the general scenario, hence the next best reasonable solution concept - Bayesian implementation, is considered.

In a Bayesian set up, agents have a prior distribution on the type profile. For agent i , prior is $p_i \in \Delta(\Theta)$. For basic regularity assume that the prior gives non-zero probability on all points of Θ (this is only a technical condition and the ensuing results can be proved without it as well). These priors are assumed to be common knowledge between agents and designer - hence there is no need to introduce second order beliefs over the priors and so on.

The mechanism used here is a direct mechanism with allocation function $\hat{x}_\epsilon(\cdot)$ (same as before). Given the allocation and tax functions $(\hat{x}_\epsilon, t) : \Theta \rightarrow \mathcal{X} \times \mathbb{R}^N$, the utility function in the Bayesian set up for strategy profile $\{\sigma_j : \Theta_j \rightarrow \Theta_j\}_{j \in \mathcal{N}}$ is given by, $\forall i \in \mathcal{N}$,

$$\tilde{u}_i(\sigma | \theta_i) = \mathbb{E}_{p_i(\cdot | \theta_i)} [v(\hat{x}_\epsilon(\sigma(\theta)); \theta_i) - t_i(\sigma(\theta))]. \quad (8)$$

where θ_i is the true type of agent i . The Bayesian implementation condition—also known as Bayesian Strategy Incentive Compatibility (BSIC)—for the direct mechanism is that the truthful strategy $\sigma_i^*(\theta_i) = \theta_i, \forall \theta_i \in \Theta_i, \forall i \in \mathcal{N}$ must be a Bayesian Nash equilibrium (BNE)³ for the induced Bayesian game.

In addition, it is required the tax function to have the Strong Budget Balance (SBB) property, i.e.,

$$\sum_{i \in \mathcal{N}} t_i(\psi) = 0, \quad \forall \psi \in \Theta. \quad (9)$$

We restrict attention to optimization (1) and priors $\{p_i(\cdot)\}_{i \in \mathcal{N}}$ that satisfy the following conditions.

³Strategy $\sigma^* = (\sigma_i^* : \Theta_i \rightarrow \Theta_i)_{i \in \mathcal{N}}$ is a BNE if $\forall i \in \mathcal{N}, \forall \theta_i \in \Theta_i, \forall \sigma_i^* : \Theta_i \rightarrow \Theta_i; \tilde{u}_i(\sigma_i^*, \sigma_{-i}^* | \theta_i) \geq \tilde{u}_i(\sigma_i', \sigma_{-i}^* | \theta_i)$.

Condition (A_B): Assume that $\exists \epsilon_{max} > 0$ such that $H(\epsilon) > 0$ for all $0 \leq \epsilon < \epsilon_{max}$, where

$$H(\epsilon) := \min_{i \in \mathcal{N}} \min_{(\theta_i, \phi_i) \in \mathcal{F}_i} \mathbb{E}_{p_i(\cdot | \theta_i)} [h_i(\theta_i, \phi_i, \theta_{-i}, \epsilon)], \quad (10a)$$

$$h_i(\theta_i, \phi_i, \theta_{-i}, \epsilon) := \sum_{j \in \mathcal{N}} \left(v(\hat{x}_\epsilon(\theta_i, \theta_{-i}); \theta_j) - v(\hat{x}_\epsilon(\phi_i, \theta_{-i}); \theta_j) \right). \quad (10b)$$

This is the most general form of the assumption needed about the optimization here; after proving Theorem IV.1, corollaries are stated with stricter assumptions than above.

Condition (B): Assume that for any non-zero vector $R := (R(\psi))_{\psi \in \Theta}$ with $R(\psi) \in \mathbb{R}$, there does not exist any $\lambda := (\lambda_k(\theta_k, \phi_k))_{k \in \mathcal{N}, (\theta_k, \phi_k) \in \mathcal{F}_k}$ with $\lambda_k(\theta_k, \phi_k) \in \mathbb{R}_+$ such that $\forall i \in \mathcal{N}, \forall \psi \in \Theta$,

$$p_i(\psi_{-i} | \psi_i) \sum_{\substack{\phi_i \in \Theta_i \\ \phi_i \neq \psi_i}} \lambda_i(\phi_i, \psi_i) - \sum_{\substack{\theta_i \in \Theta_i \\ \theta_i \neq \psi_i}} p_i(\psi_{-i} | \theta_i) \lambda_i(\theta_i, \psi_i) = R(\psi). \quad (11)$$

This condition was first introduced in [14] and subsumes the case of conditionally independent priors, i.e., $p_j(\psi_{-j} | \psi_j) = p_j(\psi_{-j})$ (refer to [14] for an example of priors which are not conditionally independent but still satisfy the condition above).

The second contribution of this paper is summarized in the following Theorem.

Theorem IV.1. *If Conditions (A_B) and (B) are satisfied then for all $0 \leq \epsilon < \epsilon_{max}$, there exist taxes $(t_i(\phi))_{\phi \in \Theta_i, i \in \mathcal{N}}$ that satisfy BSIC (which implies implementation in BNE) and SBB.*

Proof: Please see Appendix C. ■

Stated below are corollaries which successively use stricter conditions.

Corollary IV.2. *If $\exists \epsilon_{max} > 0$ such that for all $0 \leq \epsilon < \epsilon_{max}$,*

$$\min_{i \in \mathcal{N}} \min_{(\theta_i, \phi_i) \in \mathcal{F}_i} \min_{\theta_{-i} \in \Theta_{-i}} h_i(\theta_i, \phi_i, \theta_{-i}, \epsilon) > 0 \quad (12)$$

and Condition (B) is satisfied then for all $0 \leq \epsilon < \epsilon_{max}$ there exist taxes which satisfy BSIC and SBB.

Corollary IV.3. *If $\exists \epsilon_{max} > 0$ such that for all $0 \leq \epsilon < \epsilon_{max}$, and $\forall i \in \mathcal{N}, (\theta_i, \phi_i) \in \mathcal{F}_i, \theta_{-i} \in \Theta_{-i}$,*

$$K(\epsilon) + \epsilon g_i(\theta_i, \phi_i, \theta_{-i}, \epsilon) > 0, \quad (13a)$$

$$\text{where } K(\epsilon) := \min_{i \in \mathcal{N}} \min_{(\theta_i, \phi_i) \in \mathcal{F}_i} \min_{\theta_{-i} \in \Theta_{-i}} K_i(\theta_i, \phi_i, \theta_{-i}, \epsilon), \quad (13b)$$

$$g_i(\theta_i, \phi_i, \theta_{-i}, \epsilon) := \sum_{j \in \mathcal{N}} f(v(\hat{x}_\epsilon(\theta_i, \theta_{-i}); \theta_j)) - f(v(\hat{x}_\epsilon(\phi_i, \theta_{-i}); \theta_j)) \quad (13c)$$

and Condition (B) is satisfied then for all $0 \leq \epsilon < \epsilon_{max}$ there exist taxes which satisfy BSIC and SBB.

Since (13a) \Rightarrow (12) \Rightarrow Condition (A_B), hence these Corollaries follow from Theorem IV.1.

V. FULL BAYESIAN IMPLEMENTATION

In accordance with the majority of the literature on Bayesian implementation (see for e.g. [15], [16], [3], [17]), the main result in the previous section aimed at BSIC condition for implementation. This ensures that truth-telling is a BNE but gives no information about other possible BNE. In general this might be a problem, since when the Bayesian game is played, the designer cannot predict in advance which BNE will be achieved. The justification used in such a situation is that of “focusing”⁴. Along with the mechanism, the selected BNE is also announced by the designer. All the agents are then focused towards this particular BNE and in anticipation that others will be playing according to it, they too choose to play it.

In this section we modify our mechanism by augmenting each agents’ message space to include one additional message - which takes values in a continuous space - with the specific aim of achieving full implementation i.e. truth-telling as the only BNE. This will add to the robustness of the mechanism and will put to rest any equilibrium selection issues. Such a modification only requires that the BSIC constraints (see (30)) can be satisfied with strict inequality. This gives the designer room to alter taxes.

For general Bayesian mechanism design, authors in [18] have derived sufficient conditions under which augmenting the message space results in full implementation. However a specific mechanism is not derived.

To keep analysis straightforward we make the next assumptions.

Assumption V.1. Assume private consumption i.e. any agent i ’s utility is affected only by level of his/her consumption x_i .

Assumption V.2. Assume that $v(x_i, \theta_i)$ is strictly concave and is differentiable w.r.t. x_i at all values of x_i, θ_i .

Extending to general public goods is possible, but will make the proofs more technical and obfuscate the basic idea behind them. Differentiability of the utility function is essential to get an explicit modification.

The new message space, allocation and taxes for any agent i are

$$\mathcal{M}_i = \Theta_i \times [-\delta, +\delta], \quad m_i = (\phi_i, y_i). \quad (14a)$$

$$\hat{x}_{\epsilon, i}(m) = \hat{x}_{\epsilon, i}(\phi) + y_i, \quad (14b)$$

$$\tilde{t}_i(m) = \begin{cases} \hat{t}_i(\phi) + y_i v'(\hat{x}_{\epsilon, i}(\phi); \phi_i) & \text{if } y \in (-\delta, +\delta), \\ B & \text{if } y = \pm\delta. \end{cases} \quad (14c)$$

where $\hat{x}_\epsilon(\phi)$ is the allocation from (1) and $\hat{t}_i(\phi)$ is a tax designed for BSIC and SBB (from the previous section). Also, $\delta, B > 0$ are constants chosen by the designer. Here δ will be a small constant and how it is chosen will be clear in the proof of Theorem V.3. The constant B is chosen to be large enough so that no rational agent will ever choose message $y_i = \pm\delta$. This is possible because due to the type sets being discrete, the utilities are bounded. The modification above

⁴This is the same justification as the one used for working without loss of generality, with the Revelation principle and direct mechanisms.

allows agents to change their allocation by a small amount $y_i \in [-\delta, +\delta]$. For this increase/decrease in allocation they are charged/subsidized at the “market” price $v'(\hat{x}_{\epsilon,i}(\phi); \phi_i)$ corresponding to the truth-telling strategy. This modification serves the purpose of giving agents more delicate control of their allocation and utility than in a discrete set-up. Technically, this allows the designer to disrupt any BNE at which the price an agent is capable of paying (i.e. his/her derivative without tax terms) doesn’t match the market price designed for truthful strategies.

Theorem V.3. *Assuming there exists taxes that satisfy BSIC constraints (see for instance (30a)) with strict inequality, for the mechanism defined in (14) at any true type profile $\theta = (\theta_i)_{i \in \mathcal{N}}$,*

(1) *message $m^* = (\theta, \underline{0})$ is a BNE.*

Furthermore, assuming $\forall i \in \mathcal{N}$, $(\theta_i, \phi_i) \in \mathcal{F}_i$, the sign of the quantity $v'(\hat{x}_{\epsilon,i}(\phi_i, \phi_{-i}); \theta_i) - v'(\hat{x}_{\epsilon,i}(\phi_i, \phi_{-i}); \phi_i)$ remains same $\forall \phi_{-i} \in \Theta_{-i}$,

(2) *any message other than m^* isn’t a BNE.*

Hence the mechanism in (14) achieves full Bayesian implementation.

Proof: Please see Appendices E and F. ■

The no-crossing condition above gives a more definite meaning to private types. In SoU or related problems like here (1), the level of allocation is determined more by the rate of growth of utilities than the value. This is because at optimum, trade-offs between giving infinitesimal allocation to one agent vs. another will be determined by the slope of their utilities at those points. Thus by the above condition, one can implicitly order types - from the one providing smallest slope to the highest.

A stronger assumption (than above) that works is that the sign of the quantity $v'(x_i; \theta_i) - v'(x_i; \phi_i)$ is same $\forall x_i \in \text{Proj.}(\mathcal{X})$ (projection onto the i -th dimension). This might be useful in some cases since verifying this only requires utility functions and type sets whereas ϵ and the optimization solution aren’t required.

The above assumptions are in fact two of many that can be made to get the result. The proof of Part (2) of Theorem V.3 requires the quantity

$$\sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) \left[v'(\hat{x}_i(\sigma(\theta)); \theta_i) - v'(\hat{x}_i(\sigma(\theta)); \sigma_i(\theta_i)) \right] \quad (15)$$

to remain non-zero whenever $\sigma_i(\theta_i) \neq \theta_i$. In general, any condition that ensures that for any non-truthful strategy σ there exist an agent i and type θ_i for which the above expression is non-zero, would do as well. For example, given type sets, utilities and the optimization, a condition on the priors might thus be sufficient. Furthermore, such condition wouldn’t significantly reduce the set of possible priors since it only requires few equations to remain non-zero.

Finally, the mechanism presented here has SBB property only at BNE. However, as in proof of Theorem IV.1, using the d’AGV form for budget balanced taxes i.e. $\tilde{t}_i(m) = \tilde{z}_i(m) - \frac{1}{N-1} \sum_{j \neq i} \tilde{z}_j(m)$ a straightforward modification can

make the mechanism in (14) budget balanced off-equilibrium as well.

VI. NUMERICAL RESULTS / EXAMPLES

This section contains numerical examples which have been evaluated to ascertain the scope of application of the existence results provided in the previous sections. In particular we are interested in evaluating the gain in overall fairness w.r.t. allocation and taxes, attained with the proposed method.

Consider $\Theta_i = S \triangleq \{\theta^H, \theta^L\} \forall i \in \mathcal{N}$, where $\theta^H > \theta^L$ and utilities are quadratic with private consumption i.e. $\forall i \in \mathcal{N}, \forall \theta_i \in S, \forall x \in \mathbb{R}_+^N$,

$$v(x; \theta_i) := 2\theta_i x_i - \theta_i x_i^2. \quad (16)$$

The constraint set is $\mathcal{X} = \{x \in \mathbb{R}_+^N \mid \sum_i x_i = 1\}$ and $g_\epsilon(z) = z - \epsilon z^2$. Thus the Centralized optimization problem is

$$\hat{x}_\epsilon(\theta) = \arg \max_{x \in \mathcal{X}} \sum_{i \in \mathcal{N}} (2\theta_i x_i - \theta_i x_i^2) - \epsilon (2\theta_i x_i - \theta_i x_i^2)^2. \quad (17)$$

We consider the well known Gini coefficient (GC)⁵ as a measure of disparity in allocation. At $\epsilon = 0$ the objective is exactly the SoU. As ϵ starts increasing from 0 onwards, the optimization problem is transformed such that higher utilities will be weighed less than lower ones - thereby giving closer to equal distribution of allocation.

For the numerical analysis we consider Bayesian implementation (although Dominant implementation is also possible in this two-type set up) with $\theta^H = 1$, $\theta^L = 0.75$ and two cases $N = 10, 90$ and vary ϵ . Figures 1, 2 depict the GC of the utility at optimal allocation, $\{v(\hat{x}_{\epsilon,i}(\theta), \theta_i)\}_{i \in \mathcal{N}}$ at $N = 10$ and $N = 90$, respectively. This is done for various type profiles, which due to symmetry can be defined by the number, $m \in \{0, 1, \dots, N\}$, of agents with type θ^H . Also included in the plots is the mean GC where m is chosen with Binomial($N, 0.5$) distribution for $N = 10$ and Binomial($N, 0.1$) for $N = 90$. Note that in both cases there is a potential for large gains in fairness, as long as a large enough ϵ exists, at which implementation is possible.

In general there are two sources of upper bound on ϵ . One is through the ϵ_{max} defined in Condition (A_B) and other is through the well-defined-ness of the optimization problem (17). For $\epsilon < \frac{1}{2\theta^H} = 0.5$, optimization (17) is a convex optimization problem and hence has a unique optimizer. One can verify however that the optimization continues to have a unique optimizer even beyond 0.5 and for all values of ϵ within the range depicted in Fig. 1, 2. Therefore we only need to look at Condition (A_B). For $N = 10$, Condition (A_B) gives $\epsilon_{max} \approx 1.504$ and for $N = 90$ we get $\epsilon_{max} > 5$. With this we limit our plots to $\epsilon \leq 1.5$ for $N = 10$ and $\epsilon \leq 5$ for $N = 90$.

⁵Defined as the ratio of mean of the difference between every possible pair of data points with mean size. The lower the value of GC, the more equitable the allocation. GC of 0 is perfectly fair and GC of 1 is absolutely unfair - everyone except an individual receiving 0. GC is independent of scale, hence can also be used in comparing fairness across different settings.

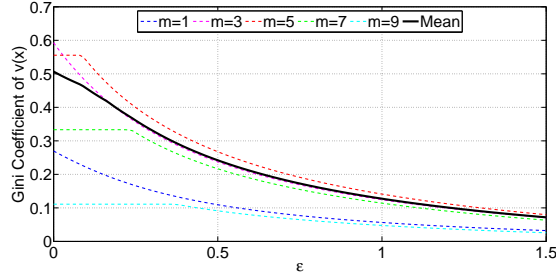


Fig. 1. Gini Coefficient of $\{v(\hat{x}_\epsilon(\theta); \theta_i)\}_{i \in \mathcal{N}}$ vs ϵ , $N = 10$.

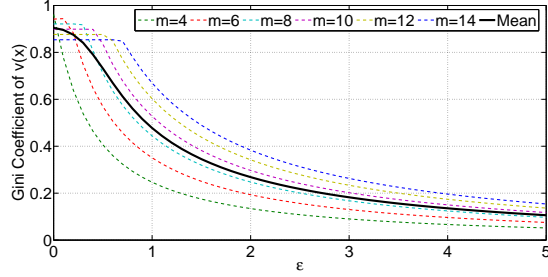


Fig. 2. Gini Coefficient of $\{v(\hat{x}_\epsilon(\theta); \theta_i)\}_{i \in \mathcal{N}}$ vs ϵ , $N = 90$.

In the figures we see that the mean GC can be reduced by 42 and 80 percentage points for $N = 10$ and $N = 90$, respectively.

Finally, Figures 3, 4 depict the standard deviation in the tax vector $t(\theta)$ for various type profiles as well as their mean. Here in each case the tax is chosen such that it minimizes the average variance within the feasible space of taxes, as dictated by the BSIC constraints. As ϵ increases, for $N = 10$ the standard deviation in taxes paid can be driven from 0.134 to 0.0001 over the range of permissible ϵ . As mentioned in the Introduction, this means that tax fluctuation between various agents can be made lower, which can prevent a situation where taxes required to be paid from agents are not within their means. For $N = 90$, standard deviation reduces by more than one order of magnitude, as it goes from 0.046 to 0.0035.

VII. DISCUSSION AND CONCLUSIONS

This paper introduces a concept of fairness in resource allocation problems pertaining to SoU maximization. The fairness aspect is adjustable (through the selection of the

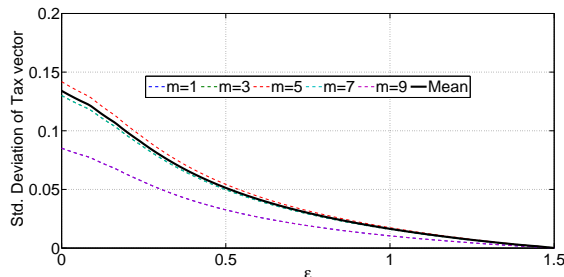


Fig. 3. Standard Deviation of $t(\theta)$ vs ϵ , $N = 10$.

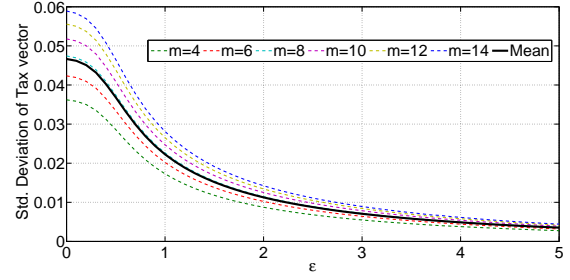


Fig. 4. Standard Deviation of $t(\theta)$ vs ϵ , $N = 90$.

parameter ϵ and function $f(\cdot)$) in a family of functions, thus giving a wide variety of criteria that a designer may choose at their own discretion. The main result in this paper is the proof of existence of mechanisms that implement the fairer allocation in Dominant and Bayesian equilibria (in respective cases). Numerical results indicate that through the proposed techniques there are significant gains in fairness of allocation, within the permissible limits of the design method.

Although the form considered here is $g_\epsilon(z) = z - \epsilon f(z)$, it is easy to see that the results can be extended to cases where f depends on ϵ ; as long as terms of the form $\epsilon \cdot f$ go to 0 as $\epsilon \rightarrow 0$. Ideally one would like to consider the class of $g_\epsilon(z) = z^{1-\epsilon}$, so as to reconcile with known fair social utilities such as the geometric mean and min utility. This however may not be a practical necessity since as indicated by the results in Section VI, even the form $g_\epsilon(z) = z - \epsilon z^2$ provides a significant reduction in GC, as well as the standard deviation of taxes. Also, as demonstrated in the Introduction, even theoretically the form $g_\epsilon(z) = z - \epsilon f(z)$ can be considered as a close approximation to many other families (including the one above) using the Taylor's series.

Finally a modification of our mechanism was presented, which guarantees truth-telling as the only BNE. This was done by adding one continuous message per agent, other than his/her type. This modification is especially useful in situations where selection of equilibria is too complex to predict.

REFERENCES

- [1] L. Hurwicz and S. Reiter, *Designing economic mechanisms*. Cambridge University Press, 2006.
- [2] E. Maskin and T. Sjöström, "Implementation theory," *Handbook of social Choice and Welfare*, vol. 1, pp. 237–288, 2002.
- [3] M. O. Jackson, "A crash course in implementation theory," *Social choice and welfare*, vol. 18, no. 4, pp. 655–708, 2001.
- [4] V. Krishna, *Auction theory*. Academic press, 2009.
- [5] T. Groves, "Incentives in teams," *Econometrica: Journal of the Econometric Society*, pp. 617–631, 1973.
- [6] E. Wei, A. Malekian, and A. Ozdaglar, "Competitive equilibrium in electricity markets with heterogeneous users and price fluctuation penalty," *LIDS report*, vol. 2918, 2014.
- [7] R. Cole, V. Gkatzelis, and G. Goel, "Mechanism design for fair division: Allocating divisible items without payments," in *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, ser. EC '13. New York, NY, USA: ACM, 2013, pp. 251–268. [Online]. Available: <http://doi.acm.org/10.1145/2482540.2482582>
- [8] M. M. Pai and R. Vohra, "Auction design with fairness concerns: Subsidies vs. set-asides," Discussion Paper, Center for Mathematical Studies in Economics and Management Science, Tech. Rep., 2012.

- [9] H. R. Varian, "Equity, envy, and efficiency," *Journal of economic theory*, vol. 9, no. 1, pp. 63–91, 1974.
- [10] E. Maskin and J. Riley, "Optimal auctions with risk averse buyers," *Econometrica: Journal of the Econometric Society*, pp. 1473–1518, 1984.
- [11] S. Dughmi and Y. Peres, "Mechanisms for risk averse agents, without loss," CoRR, vol. abs/1206.2957, 2012. [Online]. Available: <http://arxiv.org/abs/1206.2957>
- [12] C. d'Aspremont, J. Cremer, and L.-A. Gérard-Varet, "Incentives and the existence of pareto-optimal revelation mechanisms," *Journal of Economic Theory*, vol. 51, no. 2, pp. 233–254, 1990.
- [13] C. d'Aspremont, J. Crémer, and L.-A. Gérard-Varet, "Balanced bayesian mechanisms," *Journal of Economic Theory*, vol. 115, no. 2, pp. 385–396, 2004.
- [14] C. d'Aspremont and L.-A. Gérard-Varet, "Incentives and incomplete information," *Journal of Public economics*, vol. 11, no. 1, pp. 25–45, 1979.
- [15] T. Borgers, D. Krahmer, and R. Strausz, *An introduction to the theory of mechanism design*. Oxford University Press, 2015.
- [16] M. O. Jackson, "Bayesian implementation," *Econometrica: Journal of the Econometric Society*, pp. 461–477, 1991.
- [17] J. D. Hartline, "Bayesian mechanism design," *Theoretical Computer Science*, vol. 8, no. 3, pp. 143–263, 2012.
- [18] D. Mookherjee and S. Reichelstein, "Implementation via augmented revelation mechanisms," *The Review of Economic Studies*, pp. 453–475, 1990.
- [19] R. T. Rockafellar, *Convex analysis*. Princeton university press, 1970, no. 28.

APPENDIX A

PROOF OF THEOREM III.1

Proof: Note that the taxes are a finite collection of variables, since Θ, \mathcal{N} are both finite sets. For the DSIC constraints to be satisfied, the following constraints must hold $\forall i \in \mathcal{N}, \forall \theta_{-i} \in \Theta_{-i}$,

$$\begin{aligned} v(\hat{x}_\epsilon(\theta_i^H, \theta_{-i}); \theta_i^H) - t_i(\theta_i^H, \theta_{-i}) \\ \geq v(\hat{x}_\epsilon(\theta_i^L, \theta_{-i}); \theta_i^H) - t_i(\theta_i^L, \theta_{-i}), \end{aligned} \quad (18a)$$

$$\begin{aligned} v(\hat{x}_\epsilon(\theta_i^L, \theta_{-i}); \theta_i^L) - t_i(\theta_i^L, \theta_{-i}) \\ \geq v(\hat{x}_\epsilon(\theta_i^H, \theta_{-i}); \theta_i^L) - t_i(\theta_i^H, \theta_{-i}). \end{aligned} \quad (18b)$$

This gives truth-telling as a dominant strategy for agent i regardless of types of others. For IR, the following constraints must be satisfied $\forall i \in \mathcal{N}, \forall \theta_{-i} \in \Theta_{-i}$,

$$t_i(\theta_i^H, \theta_{-i}) \leq v(\hat{x}_\epsilon(\theta_i^H, \theta_{-i}); \theta_i^H), \quad (19a)$$

$$t_i(\theta_i^L, \theta_{-i}) \leq v(\hat{x}_\epsilon(\theta_i^L, \theta_{-i}); \theta_i^L). \quad (19b)$$

From the above sets of constraints, it is clear that one can design $(t_i(\theta_i^H, \theta_{-i}), t_i(\theta_i^L, \theta_{-i}))$ separately for each $i \in \mathcal{N}, \theta_{-i} \in \Theta_{-i}$. So for any i, θ_{-i} , the constraints can be rewritten in the form

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \leq \begin{bmatrix} A_{HH} - A_{LH} \\ A_{LL} - A_{HL} \\ A_{HH} \\ A_{LL} \end{bmatrix} \quad (20)$$

where $(t_1, t_2) = (t_i(\theta_i^H, \theta_{-i}), t_i(\theta_i^L, \theta_{-i}))$ and

$$\begin{aligned} A_{HH} &= v(\hat{x}_\epsilon(\theta_i^H, \theta_{-i}); \theta_i^H), \quad A_{LL} = v(\hat{x}_\epsilon(\theta_i^L, \theta_{-i}); \theta_i^L), \\ A_{HL} &= v(\hat{x}_\epsilon(\theta_i^H, \theta_{-i}); \theta_i^L), \quad A_{LH} = v(\hat{x}_\epsilon(\theta_i^L, \theta_{-i}); \theta_i^H). \end{aligned} \quad (21)$$

Using the Farkas Lemma, the above system is feasible in t if and only if $\forall \lambda \in \mathbb{R}_+^4$

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}^\top \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} A_{HH} - A_{LH} \\ A_{LL} - A_{HL} \\ A_{HH} \\ A_{LL} \end{bmatrix}^\top \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} \geq 0. \quad (22)$$

The equality constraints on λ give that

$$\begin{aligned} \lambda_1 + \lambda_3 &= \lambda_2, \quad \lambda_2 + \lambda_4 = \lambda_1 \\ \Rightarrow \lambda_3 + \lambda_4 &= 0 \Rightarrow \lambda_3 = \lambda_4 = 0. \end{aligned} \quad (23)$$

So $\lambda \in \mathbb{R}_+^4$ can be parametrized as $\lambda = (\xi, \xi, 0, 0)^\top$ for $\xi \in \mathbb{R}_+$. Thus for feasibility, using Farkas Lemma, the condition that must be satisfied is

$$\begin{aligned} (A_{HH} - A_{LH}, A_{LL} - A_{HL}, A_{HH}, A_{LL}) \cdot (\xi, \xi, 0, 0) &\geq 0 \\ \Leftrightarrow \xi(A_{HH} + A_{LL} - A_{HL} - A_{LH}) &\geq 0 \\ \Leftrightarrow A_{HH} + A_{LL} - A_{HL} - A_{LH} &\geq 0. \end{aligned} \quad (24)$$

■

APPENDIX B

PROOF OF THEOREM III.2

Proof: From the definition in (2), we have

$$\begin{aligned} v(\hat{x}_\epsilon(\theta_i, \theta_{-i}); \theta_i) - \epsilon f(v(\hat{x}_\epsilon(\theta_i, \theta_{-i}); \theta_i)) \\ + \sum_{j \neq i} v(\hat{x}_\epsilon(\theta_j, \theta_{-i}); \theta_j) - \epsilon f(v(\hat{x}_\epsilon(\theta_j, \theta_{-i}); \theta_j)) \\ = v(\hat{x}_\epsilon(\phi_i, \theta_{-i}); \theta_i) - \epsilon f(v(\hat{x}_\epsilon(\phi_i, \theta_{-i}); \theta_i)) \\ + \sum_{j \neq i} v(\hat{x}_\epsilon(\phi_j, \theta_{-i}); \theta_j) - \epsilon f(v(\hat{x}_\epsilon(\phi_j, \theta_{-i}); \theta_j)) \\ + K_i(\theta_i, \phi_i, \theta_{-i}, \epsilon). \end{aligned} \quad (25)$$

Using the above twice, first with $(\theta_i, \phi_i) = (\theta^H, \theta^L), \theta_{-i} = \theta_{-i}$ and then with $(\theta_i, \phi_i) = (\theta^L, \theta^H), \theta_{-i} = \theta_{-i}$, and adding the two results in (using the notation from proof of Theorem III.2)

$$\begin{aligned} A_{HH} - \epsilon f(A_{HH}) + A_{LL} - \epsilon f(A_{LL}) \\ = K_i(\theta^H, \theta^L, \theta_{-i}, \epsilon) + K_i(\theta^L, \theta^H, \theta_{-i}, \epsilon) \\ + A_{HL} - \epsilon f(A_{HL}) + A_{LH} - \epsilon f(A_{LH}). \end{aligned} \quad (26)$$

This can be rewritten as

$$\begin{aligned} \Leftrightarrow A_{HH} + A_{LL} - A_{HL} - A_{LH} \\ = K_i(\theta^H, \theta^L, \theta_{-i}, \epsilon) + K_i(\theta^L, \theta^H, \theta_{-i}, \epsilon) \\ + \epsilon(f(A_{HH}) + f(A_{LL}) - f(A_{HL}) - f(A_{LH})) \end{aligned} \quad (27)$$

Thus it is sufficient to prove that the RHS above is non-negative. Owing to Condition (A_D), $\exists \epsilon_{max} > 0$ such that the sum of the first two terms in RHS is strictly positive for all $0 \leq \epsilon < \epsilon_{max}$ and clearly the second term can be made arbitrarily small in magnitude (by choosing a smaller ϵ_{max}).

Hence the condition in (5) is satisfied⁶ for all $0 \leq \epsilon < \tilde{\epsilon}_{max}$. Here $\tilde{\epsilon}_{max}$ is bigger or smaller than ϵ_{max} depending on whether $f(A_{HH}) + f(A_{LL}) - f(A_{HL}) - f(A_{LH})$ is positive or negative in the range $(0, \epsilon_{max})$. ■

APPENDIX C PROOF OF THEOREM IV.1

Proof: The utility for any agent i when other agents are truth-telling is

$$\tilde{u}_i(\phi_i | \theta_i) = \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) \cdot \left[v(\hat{x}_\epsilon(\phi_i, \theta_{-i}); \theta_i) - t_i(\phi_i, \theta_{-i}) \right] \quad (28)$$

where agent i 's true and quoted types are $\theta_i, \phi_i \in \Theta_i$, respectively. Here we consider taxes in the d'AGV form

$$t_i(\psi) = z_i(\psi) - \frac{1}{N-1} \sum_{\substack{j \in \mathcal{N} \\ j \neq i}} z_j(\psi) \quad \forall i \in \mathcal{N}, \psi \in \Theta. \quad (29)$$

Taxes in this form always satisfy SBB and any tax function which satisfies SBB can be written in this form. Therefore WLOG, the design variables from here onwards will be $\{z_j(\psi)\}_{j \in \mathcal{N}, \psi \in \Theta}$.

BSIC constraints can be written as: $\forall i \in \mathcal{N}, (\theta_i, \phi_i) \in \mathcal{F}_i$,

$$\tilde{u}_i(\theta_i | \theta_i) \geq \tilde{u}_i(\phi_i | \theta_i). \quad (30a)$$

$$\begin{aligned} &\Leftrightarrow \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) \left[z_i(\theta_i, \theta_{-i}) - \frac{1}{N-1} \sum_{j \neq i} z_j(\theta_i, \theta_{-i}) \right. \\ &\quad \left. - z_i(\phi_i, \theta_{-i}) + \frac{1}{N-1} \sum_{j \neq i} z_j(\phi_i, \theta_{-i}) \right] \\ &\leq \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) [v(\hat{x}_\epsilon(\theta_i, \theta_{-i}); \theta_i) - v(\hat{x}_\epsilon(\phi_i, \theta_{-i}); \theta_i)]. \end{aligned} \quad (30b)$$

So the condition for design of taxes is a linear system of inequalities and can be written in the form $Az \leq b$ where the indexing is as follows

$$z = (z_j(\psi)) \in \mathbb{R}^{D_1}, \quad b = (b(i, \theta_i, \phi_i)) \in \mathbb{R}^{D_2}, \\ A = (A(i, \theta_i, \phi_i | j, \psi)) \in \mathbb{R}^{D_2 \times D_1}. \quad (31)$$

for $D_1 = N \cdot \prod_{j=1}^N L_j$ and $D_2 = \prod_{i=1}^N (L_i^2 - L_i)$. The rest of the proof will be to show that this linear system is feasible in variable z , using the Farkas Lemma⁷ [19, pg.201].

⁶Overall the behaviour of $K_i(\theta_i, \phi_i, \theta_{-i}, \epsilon)$ w.r.t. ϵ will dictate the value of ϵ_{max} , $\tilde{\epsilon}_{max}$. This in turn will effect the usefulness of this method, since a designer might want to ensure certain minimum gains in fairness for which he/she might want to choose ϵ as large as possible.

⁷ Relevant version of the Farkas alternative result: The system

$$x \in \mathbb{R}^N, \quad Ax \leq b; \quad A \in \mathbb{R}^{M \times N}, b \in \mathbb{R}^M \quad (32)$$

is feasible iff all solutions $\lambda \in \mathbb{R}_+^M$ of $A^\top \lambda = 0$ satisfy $b^\top \lambda \geq 0$.

Consider any $\lambda = (\lambda_i(\theta_i, \phi_i)) \in \Lambda \triangleq \{\lambda_k(\theta_k, \phi_k) \in \mathbb{R}_+ \mid k \in \mathcal{N}, (\theta_k, \phi_k) \in \mathcal{F}_k\}$ that satisfies $A^\top \lambda = 0$, i.e. $\forall j \in \mathcal{N}, \psi \in \Theta$,

$$\begin{aligned} &\sum_{i \in \mathcal{N}} \sum_{(\theta_i, \phi_i) \in \mathcal{F}_i} A(i, \theta_i, \phi_i | j, \psi) \lambda_i(\theta_i, \phi_i) = 0 \\ &\Leftrightarrow p_j(\psi_{-j} | \psi_j) \sum_{\phi_j \neq \psi_j} \lambda_j(\psi_j, \phi_j) \\ &\quad - \sum_{\theta_j \neq \psi_j} p_j(\psi_{-j} | \theta_j) \lambda_j(\theta_j, \psi_j) \\ &\quad - \frac{1}{N-1} \sum_{k \neq j} \left[p_k(\psi_{-k} | \psi_k) \sum_{\phi_k \neq \psi_k} \lambda_k(\psi_k, \phi_k) \right. \\ &\quad \left. - \sum_{\theta_k \neq \psi_k} p_k(\psi_{-k} | \theta_k) \lambda_k(\theta_k, \psi_k) \right] = 0. \quad (33) \end{aligned}$$

Above equation can be rearranged to give that $\forall j \in \mathcal{N}, \psi \in \Theta$,

$$\begin{aligned} &p_j(\psi_{-j} | \psi_j) \sum_{\phi_j \neq \psi_j} \lambda_j(\psi_j, \phi_j) \\ &\quad - \sum_{\theta_j \neq \psi_j} p_j(\psi_{-j} | \theta_j) \lambda_j(\theta_j, \psi_j) \\ &= \frac{1}{N} \sum_{k \in \mathcal{N}} \left[p_k(\psi_{-k} | \psi_k) \sum_{\phi_k \neq \psi_k} \lambda_k(\psi_k, \phi_k) \right. \\ &\quad \left. - \sum_{\theta_k \neq \psi_k} p_k(\psi_{-k} | \theta_k) \lambda_k(\theta_k, \psi_k) \right]. \quad (34) \end{aligned}$$

Denote the RHS above by $R(\psi)$ - note that it depends only on ψ and not j .

Lemma C.1. *If $\lambda \in \Lambda$ satisfies (34) then for any $j \in \mathcal{N}, \psi_j \in \Theta_j$,*

$$\sum_{\substack{\theta_j \in \Theta_j \\ \theta_j \neq \psi_j}} \lambda_j(\theta_j, \psi_j) = \sum_{\substack{\phi_j \in \Theta_j \\ \phi_j \neq \psi_j}} \lambda_j(\psi_j, \phi_j) \quad (35)$$

Proof: Please see Appendix D. ■

With the application of above Lemma, one can rewrite LHS of (34) to get that $\forall j, \psi$,

$$\begin{aligned} &p_j(\psi_{-j} | \psi_j) \sum_{\phi_j \neq \psi_j} \lambda_j(\phi_j, \psi_j) - \sum_{\theta_j \neq \psi_j} p_j(\psi_{-j} | \theta_j) \lambda_j(\theta_j, \psi_j) \\ &= R(\psi) \quad (36) \end{aligned}$$

Condition (B) on priors states that there exist no $\lambda \in \Lambda$ such that above holds for a non-zero R . Hence $A^\top \lambda = 0$ implies that $R \equiv 0$, therefore (by (34)) $\forall j \in \mathcal{N}, \forall \psi \in \Theta$,

$$\begin{aligned} &p_j(\psi_{-j} | \psi_j) \sum_{\phi_j \neq \psi_j} \lambda_j(\psi_j, \phi_j) \\ &= \sum_{\theta_j \neq \psi_j} p_j(\psi_{-j} | \theta_j) \lambda_j(\theta_j, \psi_j). \quad (37) \end{aligned}$$

Next we show that for all $\lambda \in \Lambda$ that satisfy (37) we have $b^\top \lambda \geq 0$ i.e.

$$\sum_{i \in \mathcal{N}} \sum_{(\theta_i, \phi_i) \in \mathcal{F}_i} b(i, \theta_i, \psi_i) \lambda_i(\theta_i, \phi_i) \geq 0 \quad (38a)$$

$$\Leftrightarrow \sum_{\substack{i \in \mathcal{N}, \\ (\theta_i, \phi_i) \in \mathcal{F}_i}} \lambda_i(\theta_i, \phi_i) \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) \cdot$$

$$\left[v(\hat{x}_\epsilon(\theta_i, \theta_{-i}); \theta_i) - v(\hat{x}_\epsilon(\phi_i, \theta_{-i}); \theta_i) \right] \geq 0. \quad (38b)$$

Proving this will finish the proof by Farkas Lemma.

For any $i, \theta_i, \phi_i, \theta_{-i}$, denote $x = \hat{x}_\epsilon(\theta_i, \theta_{-i})$, $\tilde{x} = \hat{x}_\epsilon(\phi_i, \theta_{-i})$. Rearranging terms from (2), gives

$$\begin{aligned} v(x; \theta_i) - v(\tilde{x}; \theta_i) = \\ \sum_{\substack{j \in \mathcal{N} \\ j \neq i}} \left(v(\tilde{x}; \theta_j) - v(x; \theta_j) \right) + \epsilon \sum_{j \in \mathcal{N}} \left(f(v(x; \theta_i)) - f(v(\tilde{x}; \theta_i)) \right) \\ + K_i(\theta_i, \phi_i, \theta_{-i}, \epsilon). \end{aligned} \quad (39)$$

Denote the RHS expression in (39) as $\eta = \eta_1 + \eta_2$, where

$$\begin{aligned} \eta_1 &= \sum_{\substack{j \in \mathcal{N} \\ j \neq i}} \left(v(\tilde{x}; \theta_j) - v(x; \theta_j) \right), \\ \eta_2 &= \epsilon \sum_{j \in \mathcal{N}} \left(f(v(x; \theta_j)) - f(v(\tilde{x}; \theta_j)) \right) + K_i(\theta_i, \phi_i, \theta_{-i}, \epsilon). \end{aligned} \quad (40)$$

Now continuing from LHS of (38b),

$$\text{LHS of (38b)} = \sum_{\substack{i \in \mathcal{N} \\ (\theta_i, \phi_i) \in \mathcal{F}_i}} \lambda_i(\theta_i, \phi_i) \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) (\eta_1 + \eta_2). \quad (41)$$

For any fixed i , consider the summation with only η_1 first

$$\begin{aligned} &\sum_{(\theta_i, \phi_i) \in \mathcal{F}_i} \lambda_i(\theta_i, \phi_i) \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) \cdot \\ &\quad \sum_{\substack{j \in \mathcal{N} \\ j \neq i}} \left(v(\hat{x}_\epsilon(\phi_i, \theta_{-i}); \theta_j) - v(\hat{x}_\epsilon(\theta_i, \theta_{-i}); \theta_j) \right) \\ &= \sum_{\substack{j \in \mathcal{N} \\ j \neq i}} \sum_{\substack{\phi_i \in \Theta_i, \\ \theta_{-i} \in \Theta_{-i}}} v(\hat{x}_\epsilon(\phi_i, \theta_{-i}); \theta_j) \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) \lambda_i(\theta_i, \phi_i) \\ &\quad - \sum_{\substack{j \in \mathcal{N} \\ j \neq i}} \sum_{\substack{(\theta_i, \phi_i) \in \mathcal{F}_i \\ \theta_{-i} \in \Theta_{-i}}} \lambda_i(\theta_i, \phi_i) p_i(\theta_{-i} | \theta_i) v(\hat{x}_\epsilon(\theta_i, \theta_{-i}); \theta_j). \end{aligned} \quad (42)$$

By (37), the inside summation in the first term is equal to $p_i(\theta_{-i} | \phi_i) \sum_{\psi_i} \lambda(\phi_i, \psi_i)$. Incorporating this and changing variables of summation appropriately gives the overall summation from (42) equal to 0. Now consider the term in RHS of (41) with η_2

$$\sum_{\substack{i \in \mathcal{N} \\ (\theta_i, \phi_i) \in \mathcal{F}_i}} \lambda_i(\theta_i, \phi_i) \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) \cdot \eta_2. \quad (43)$$

Rearranging terms in η_2 , we can write

$$\eta_2 = \sum_{j \in \mathcal{N}} \left(v(\hat{x}_\epsilon(\theta_i, \theta_{-i}); \theta_j) - v(\hat{x}_\epsilon(\phi_i, \theta_{-i}); \theta_j) \right). \quad (44)$$

Therefore by Condition (A_B), $\exists \epsilon_{max} > 0$ such that for all $0 \leq \epsilon < \epsilon_{max}$, the inside summation in (43) is non-negative $\forall i \in \mathcal{N}, (\theta_i, \phi_i) \in \mathcal{F}_i$. This finishes the proof by Farkas Lemma, since the expression in (38a) is now shown to be non-negative for all positive $0 \leq \epsilon < \epsilon_{max}$. ■

APPENDIX D PROOF OF LEMMA C.1

Proof:

$$\sum_{\substack{\theta_j \in \Theta_j \\ \theta_j \neq \psi_j}} \lambda_j(\theta_j, \psi_j) - \sum_{\substack{\phi_j \in \Theta_j \\ \phi_j \neq \psi_j}} \lambda_j(\psi_j, \phi_j) \quad (45a)$$

$$\begin{aligned} &= \sum_{\psi_j} \left[p_j(\psi_{-j} | \psi_j) \sum_{\theta_j \neq \psi_j} \lambda_j(\theta_j, \psi_j) \right. \\ &\quad \left. - \sum_{\phi_j \neq \psi_j} p_j(\psi_{-j} | \phi_j) \lambda_j(\psi_j, \phi_j) \right] \end{aligned} \quad (45b)$$

$$= \sum_{\psi_{-j}} R(\psi_j, \psi_{-j}) \quad (45c)$$

$$= \sum_{\psi_{-j}} \left[p_k(\psi_{-k} | \psi_k) \sum_{\theta_k \neq \psi_k} \lambda_k(\theta_k, \psi_k) \right. \quad (45d)$$

$$\left. - \sum_{\phi_k \neq \psi_k} p_k(\psi_{-k} | \phi_k) \lambda_k(\psi_k, \phi_k) \right], \quad k \neq j \\ = 0. \quad (45e)$$

Here (45c), (45d) follows by application of (34) and other equations are just by rearranging summation terms. ■

APPENDIX E PROOF OF PART (1), THEOREM V.3

Proof: For strategy

$$(\sigma, \rho) = \left\{ (\sigma_i : \Theta_i \rightarrow \Theta_i), (\rho_i : \Theta_i \rightarrow [-\delta, +\delta]) \right\}_{i \in \mathcal{N}}, \quad (46)$$

utility is

$$\begin{aligned} u_i(\sigma, \rho | \theta_i) &= \sum_{\theta_{-i}} p_i(\theta_{-i} | \theta_i) \left[v(\hat{x}_i(\sigma(\theta)) + \rho_i(\theta_i); \theta_i) \right. \\ &\quad \left. - \hat{t}_i(\sigma(\theta)) - \rho_i(\theta_i) v'(\hat{x}_{\epsilon, i}(\sigma(\theta)); \sigma_i(\theta_i)) \right] \end{aligned} \quad (47)$$

First in this proof we establish that for any agent i , true type θ_i , when other agents use strategy $m_{-i}^* = (\theta_{-i}, \underline{0})$, strategy $\sigma_i(\theta_i) = \theta_i$ gives higher utility than $\sigma_i(\theta_i) \neq \theta_i$ at any value of $\rho_i(\theta_i)$. For this start by considering the utility in (47) with $\sigma_{-i}(\theta_{-i}) = \theta_{-i}$ and $\rho_i(\theta_i) = 0$,

$$\begin{aligned} \tilde{u}_i(\sigma_i(\theta_i) | \theta_i) &= \sum_{\theta_{-i}} p_i(\theta_{-i} | \theta_i) \left[v(\hat{x}_{\epsilon, i}(\sigma_i(\theta_i), \theta_{-i}); \theta_i) \right. \\ &\quad \left. - \hat{t}_i(\sigma_i(\theta_i), \theta_{-i}) \right]. \end{aligned} \quad (48)$$

From the result in Section IV, allocation and taxes $(\hat{x}_\epsilon, \hat{t})$ satisfy BSIC condition. Hence the utility in (48) is maximized

at $\sigma_i(\theta_i) = \theta_i$. The minimum difference between the value of such an expression at $\sigma_i(\theta_i) = \theta_i$ and $\sigma_i(\theta_i) \neq \theta_i$ is strictly positive since the tax function \hat{t} satisfies BSIC with strict inequality. Now the utility u_i in (47) clearly depends continuously on $\rho_i(\theta_i)$, hence by choosing a small enough range for $\rho_i(\theta_i)$ (i.e. by choosing a small enough δ) it can be ensured that even in (47), for $\sigma_{-i}(\theta_{-i}) = \theta_{-i}$, u_i is maximized at $\sigma_i(\theta_i) = \theta_i$.

Now the only thing that remains to be proven is that when other agents quote message m_{-i}^* and $\sigma_i(\theta_i) = \theta_i$, then $\rho_i(\theta_i) = 0$ is optimal. For agent i and true type θ_i , the utility in (47) only depends on ρ_i through $\rho_i(\theta_i)$. Optimizing w.r.t. $\rho_i(\theta_i)$ (for calculating equilibrium) gives

$$\frac{\partial u_i}{\partial \rho_i(\theta_i)} = \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) \left[v'(\hat{x}_{\epsilon,i}(\theta) + \rho_i(\theta_i); \theta_i) - v'(\hat{x}_{\epsilon,i}(\theta); \sigma_i(\theta_i)) \right] = 0. \quad (49)$$

At equilibrium, the above derivative will have to be zero since the end points $\pm\delta$ cannot be maximizers - since they incur a huge tax B . Also note that since $v(x_i; \theta_i)$ is assumed to be concave in x_i (for any θ_i) the utility $u_i(\sigma, \rho | \theta_i)$ in (47) is concave in $\rho_i(\theta_i)$ and thus the condition in (49) is both necessary and sufficient for optimality w.r.t. $\rho_i(\theta_i)$.

The expression in (49) is clearly equal to 0 when $\sigma_i(\theta_i) = \theta_i$ and $\rho_i(\theta_i) = 0$. Hence applying the above result for all $\theta_i \in \Theta_i$ and $i \in \mathcal{N}$ gives that $m^* = (\theta, \underline{0})$ is a BNE. ■

APPENDIX F

PROOF OF PART (2), THEOREM V.3

Proof: Consider the derivative similar to (49) but with general strategies

$$\frac{\partial u_i}{\partial \rho_i(\theta_i)} = \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) \cdot \left[v'(\hat{x}_{\epsilon,i}(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})) + \rho_i(\theta_i); \theta_i) - v'(\hat{x}_{\epsilon,i}(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})); \sigma_i(\theta_i)) \right]. \quad (50)$$

This being zero is a necessary at BNE. Now we begin by considering the same with $\rho_i(\theta_i) = 0$,

$$\Psi = \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) \left[v'(\hat{x}_{\epsilon,i}(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})); \theta_i) - v'(\hat{x}_{\epsilon,i}(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})); \sigma_i(\theta_i)) \right]. \quad (51)$$

By the assumption in the theorem statement, for $\sigma_i(\theta_i) \neq \theta_i$, the term inside the brackets above is either positive for all θ_{-i} or negative for all θ_{-i} . Hence the expression Ψ cannot be zero for any i, θ_i when $\sigma_i(\theta_i) \neq \theta_i$. Since only discrete variables are involved in the expression Ψ , it means that all the values that Ψ can take (i.e. $\forall i, \theta_i, \sigma$ with $\sigma_i(\theta_i) \neq \theta_i$) are a discrete set (say, D) of points not containing zero.

The only difference between the expression in (50) and Ψ is the introduction of the variable $\rho_i(\theta_i) \in [-\delta, +\delta]$ - on which the expression in (50) depends continuously. This means that the set of all values (say, E) taken by $\frac{\partial u_i}{\partial \rho_i(\theta_i)}$ is the union of

continuous sets centered at the points from set D . Here the range of the continuous sets around each point in D can be continuously controlled by changing δ with the sets becoming a discrete point as $\delta \rightarrow 0$. Hence it is clear that one can find δ small enough so that even this new set E doesn't contain zero. Once such a δ is chosen it is clear that any strategy σ for which $\exists i, \theta_i$ such that $\sigma_i(\theta_i) \neq \theta_i$ cannot be a BNE since the expression in (50) being zero was a necessary condition for a BNE.

Note that each part of the proof of Theorem V.3 prescribes a positive upper limit for δ . The designer will eventually choose a δ that satisfies both.

Now consider a strategy of the form $m_i^* = (\theta_i, y_i)$ where there is truth-telling but $y_i \neq 0$. It is clear that this cannot be a BNE, since by strict concavity of $v(x_i; \theta_i)$, at $\sigma_i(\theta_i) = \theta_i$ the expression in (50) goes to zero only for $\rho_i(\theta_i) = y_i = 0$. ■